

Cross disciplinary overtures with interview data: Integrating digital practices and tools in the scholarly workflow

Stefania Scagliola

University of Luxembourg
stefania.scagliola@uni.lu

Louise Corti

University of Essex
corti@essex.ac.uk

Silvia Calamai

University of Siena
silvia.calamai@unisi.it

Norah Karrouche

Erasmus University Rotterdam
karrouche@eshcc.eur.nl

Jeannine Beeken

University of Essex
jeannine.beeken@essex.ac.uk

Arjan van Hessen

University of Twente
a.j.vanhessen@utwente.nl

Christoph Draxler

University of Muenchen
draxler@phonetik.uni-muenchen.de

Henk van den Heuvel

Radboud University
H.vandenHeuvel@let.ru.nl

Max Broekhuizen

Erasmus University Rotterdam
maksbroekhuizen@gmail.com

Khiet Truong

University of Twente
k.p.truong@utwente.nl

Abstract

There is much talk about the need for multidisciplinary approaches to research and the opportunities that have been created by digital technologies. A good example of this is the CLARIN Portal, that promotes and supports such research by offering a large suite of tools for working with textual and audio-visual data. Yet scholars who work with interview material are largely unaware of this resource and are still predominantly oriented towards familiar traditional research methods. To reach out to these scholars and assess the potential for integration of these new technologies a multidisciplinary international community of experts set out to test CLARIN-type approaches and tools on different scholars by eliciting and documenting their feedback. This was done through a series of workshops held from 2016 to 2019, and funded by CLARIN and affiliated EU funding. This paper presents the goals, the tools that were tested and the evaluation of how they were experienced. It concludes by setting out envisioned pathways for a better use of the CLARIN family of approaches and tools in the area of qualitative and oral history data analysis.

1 Introduction

Although there is much talk about the need to open up cross-disciplinary dialogue and prioritize the use of open-source software, when considering disciplines that work with interview data, we can observe a kind of pillarisation of practices. Support for the multidisciplinary approach to interview data has been endorsed by scholars such as Van den Berg et al. (2011), De Jong et al. (2014), Corti et al. (2016) and Van den Heuvel et al. (2017), but most scholars are completely unfamiliar with each other's approaches, and hesitate to take up technology. When software is used, it is often proprietary and binds scholars to a particular set of practices. This paper sets out to explore how to better exploit the rich multidisciplinary potential of interview data through the use of technology. To that end a multidisciplinary international community of experts organised a series of hands-on workshops with scholars who

This work is licenced under a Creative Commons Attribution 4.0 International License. License details: <http://creativecommons.org/licenses/by/4.0/>

Stefania Scagliola, Louise Corti, Silvia Calamai, Norah Karrouche, Jeannine Beeken, Arjan van Hessen, Cristoph Draxler, Henk van den Heuvel, Max Broekhuizen and Khiet Truong 2020. Cross disciplinary overtures with interview data: Integrating digital practices and tools in the scholarly workflow. *Selected papers from the CLARIN Annual Conference 2019*. Linköping Electronic Conference Proceedings 172: 172 126–136.

work with interview data, and tested the reception of a number of digital tools that are used at various stages of the research process. We engaged with tools for transcription, for annotation, for analysis and for emotion recognition. The workshops were held at Oxford, Utrecht, Arezzo, Munich, Utrecht and Sofia between 2016 and 2019, and were mostly sponsored by CLARIN. Participants were recruited among communities of historians, social science scholars, linguists, speech technologists, phonologists, archivists and information scientists. The website <https://oralhistory.eu/> was set up to communicate across disciplinary borders.

2 Digital tools to work with interview data

A broad diversity of practices can be observed among scholars who work with interview data. Within every discipline distinct sub-disciplines exist, and disciplinary ‘silos’ certainly complicate collaboration across computer science, humanities and social science. Scholars use the same term for very different practices, or do similar things, but give it different names (De Jong et al., 2011). Frames of interpretations differ. For instance, an oral historian will typically approach a recorded interview as an intersubjective account of a past experience, whereas another historian might consider the same source of interest only because of the factual information it conveys. A social scientist is likely to try to discover common themes and similarities and differences across a whole set of interviews, whereas a computational linguist will rely on counting frequencies and detecting collocations and co-occurrences, for similar purposes. On the other hand sociologists who interview, often seek to understand their interviewees in the same way as (oral) historians. The approaches with regard to re-use of data and anonymisation however are quite different (Van den Berg et al., 2011). The question is how they can benefit from the myriad of freely available transcription, annotation, linguistic and emotion recognition tools. To address this diversity each workshop would start off with an informative session that sketched the various ‘landscapes of practices’: the different kinds of methodological approaches to attributing meaning to interview data. This exercise in demystification was gratefully received by our audiences, as it offered the opportunity to grasp the essence of the various approaches. After this, sessions were held consisting of a short introduction followed by a step by step tutorial to practice with the various tools. This was done in groups under close supervision of experienced digital humanities scholars. After each session, a short evaluation round was held with group interviews that were recorded.

3 Creating a Transcription Tool: the T-Chain

In the first session participants had a chance to work with functionalities surrounding automatic speech recognition (ASR). Transcription is at the core of research based on interviews and scholars often require full verbatim transcripts. This often means that the focus of attention shifts from the aural dimension of the narrative to its textual representation (Portelli 2006, Boyd 2013). ASR may challenge this practice, as it could take over the laborious practice of manual transcription. Scholars are however sceptical about results of ASR, as they can be disappointing. Accepting incorrect results could however also be considered as an opportunity to rethink the standard practice of relying on full manual transcripts. By aligning audio to the ASR output researchers can easily browse through an entire interview or interview collection, making the audio more accessible and present in the process of analysis. This may lead to a practice in which only particular passages need to be fully described. With this, and other possible uses in mind, the idea of the Transcription Chain, or T-Chain was born (Van der Heuvel 2019). The first workshops were designed to collect requirements for such a tool to be used by a broad diversity of scholars.

A first version of the OH Portal was presented at the Munich workshop in September 2018 (<https://clarin.phonetik.uni-muenchen.de/apps/oh-portal/>). The workflow consists of the steps *upload*, *automatic speech recognition*, *manual correction of the ASR transcript*, *word segmentation and alignment*, and *phonetic detail* (Figure 1). The T-chain currently processes Dutch, English, Italian and German audio. The portal automatically checks the audio file format (it has to be WAV) and splits stereo recordings into separate mono audio files. These are uploaded to a server, and from there sent to different third-party providers. Thanks to a Google grant, the portal currently supports Google speech

recognisers for many languages. To aid in the selection of services, a one-line summary of the service providers' privacy policy may be displayed, together with a link to the full legal text. When the ASR results are sent back the user may choose to check and correct the transcript manually (Figure 2). For this, the transcription editor OCTRA (Pömp et al., 2017) opens within the browser.

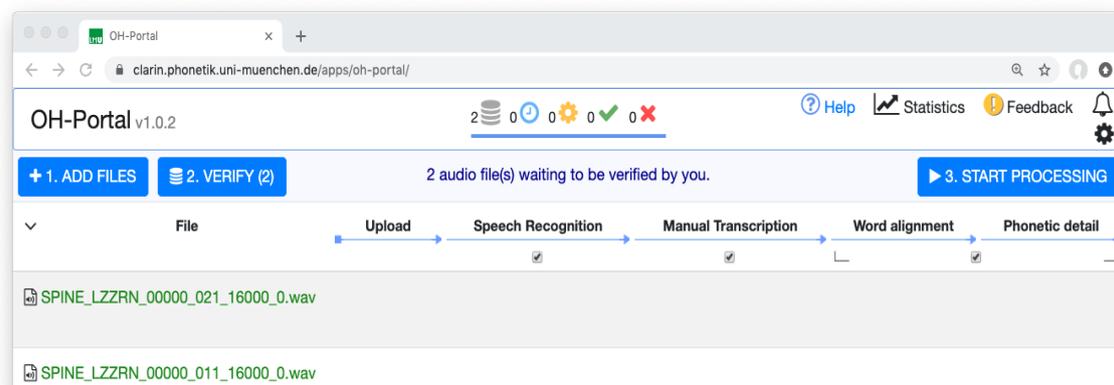


Figure 1: OH portal with two files in the workspace

This editor features a number of views and can be adapted to various transcription systems.

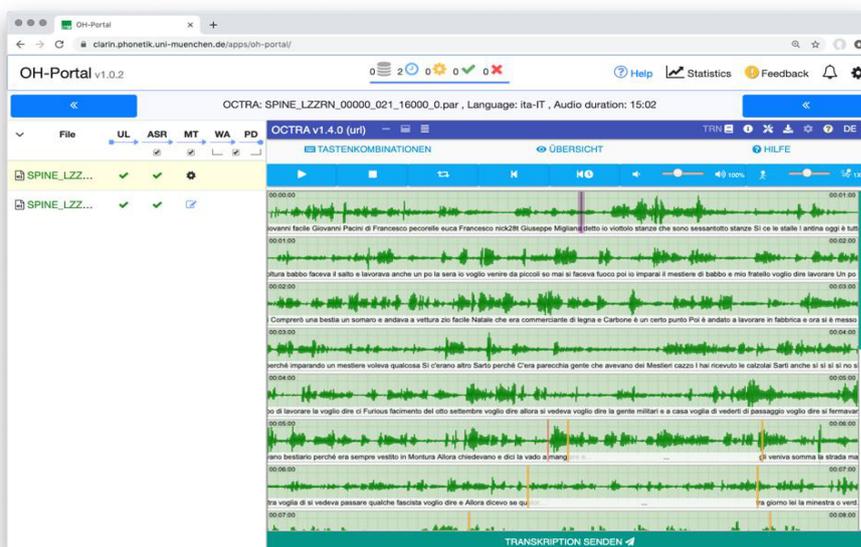


Figure 2: Transcription editor in the OH Portal window

After the manual correction of the ASR transcript, automatic word segmentation can be performed by using the WebMAUS service (Kisler et al., 2012). The result of this step is a word-based time-aligned transcript of the recording (Figure 3).

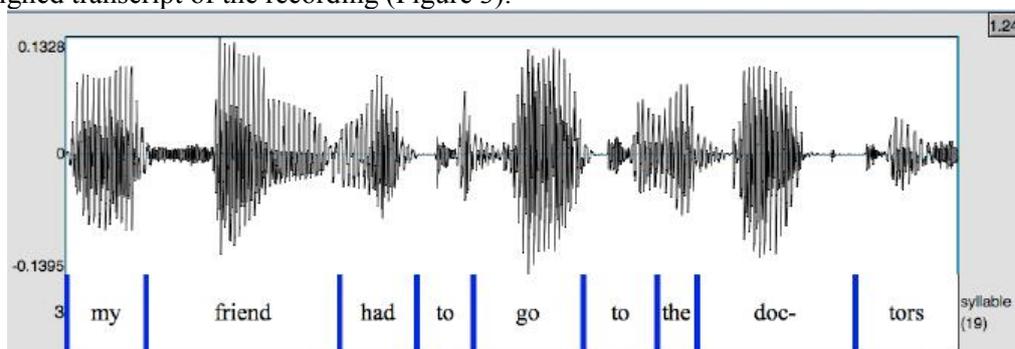


Figure 3: Example of a word alignment after ASR, visualised in Praat

At every step of the process, the output can be downloaded to the local computer in different formats, e.g. plain text, tab- or comma-separated tables, Praat TextGrid, Emu Annot-JSON, and ELAN for further processing and analysis. These exports can be made file-wise or column-wise, i.e. many files at once.

The overall performance and acceptance of the OH Portal depends on a) transcription quality, and b) on user friendliness. Transcription quality depends largely on the acoustic quality, which in turn depends on the recording situation and the speakers, and on the capabilities of the underlying ASR systems. User friendliness depends on the graphical user interface, the speed of interaction and the feeling of being in control of one's own data. We recommend that users convert their audio files using tools such as GoldWave, Audacity or To_Wave_Convertor so that they may be processed by the OH portal. The OH Portal is limited by the amount of memory available in the browser – currently, files up to 300 MB can be processed. This corresponds to approx. 118 minutes of recordings at 22.05 kHz and 16 bit mono. Note that processing long recording files means long waiting times – for a smooth operation we thus recommend that users split their recordings into short meaningful units and process them individually. At hands-on workshops, we provided sample files in different languages with durations of up to 5 minutes so that all participants could work in parallel, and experience the progress of their files through the workflow.

Finally, note that since ASR is provided by external providers, restrictions may apply. This is especially true for commercial providers who impose monthly quotas on accepted file size, recording duration or number of requests per month.

4 Using existing data annotation tools

Most researchers working with interview data use some form of annotation technique. The way these tools are used varies considerably across disciplines and the tools mostly used are proprietary. This can vary from using a pen and paper, coding of digital sources with a tool, to linguistic identification through information extraction tools.

In the area of social science, qualitative data analysis software known as QDAS is popular, and includes market leaders such as NVivo, Atlas.ti and MaxQDA. They allow analysis of text and audio-visual sources, but can be costly and tend to lock-in researchers, with no import or export capability, nor any intermediate 'portable' formats, for example, in XML (Corti and Gregory, 2011).

For the workshop we chose to offer NVivo and the open source annotation tool Elan. The first, designed for social scientists, can collect all kinds of different sources into one project, can classify and group these, mark-up text, images and audio-visual sources with thematic 'nodes', and add notes, known as 'memos'.

ELAN (<https://archive.mpi.nl/tla/elan>), in contrast, is freely available, and was developed at the Max Planck Institute for use in linguistics. Users can create 'tiers' in which annotations can be added to audio or video files, differentiating types of tiers (for example different speakers) and specifying 'parent tiers' (Wittenburg et al., 2006). The ability to annotate the audio enables users to engage with multiple dimensions of an interview from as early as the point of recording the data. This tool is also seen as very suitable to annotate audiovisual data.

4.1 Linguistic analysis tools

Text mining tools used by computational linguists could also enrich the practice of social science scholars and historians, by offering insight into the structure of language. The semantic contexts in which terms are used, are telling for how social reality is created and remembered. In non-computational methods the analysis of data starts with reading one transcript or listening to one interview at the time. Linguistic tools offer the detection of patterns in language or speech features by looking at the entire collection or to subcollections at once. This might be very useful when re-using a collection from an archive that is new to the researcher. Several features can be explored: concordances and correlations, processing syntactic tree structures, searching for named entities, and applying emotion recognition (Armaselu et al., 2019).

This part of the workshop would start with an introduction to linguistic tools and their functions: lemmatizers, syntactic parsers, named entity recognizers, auto-summarizers, tools for detecting concordances/n-grams and semantic correlations. Participants were then given a live demo of the software tools and then some step by step guided exercises with data.

The first tool to be introduced was Voyant (<https://voyant-tools.org/>), a lightweight text analysis tool that yields output on the fly (Sinclair et al., 2016). This was followed by Stanford CoreNLP (<https://stanfordnlp.github.io/CoreNLP/>), a linguistic tool that can automatically tag words in a number of different ways, such as recognizing part of speech, type of proper noun, numeric quantities, and more (Manning et al., 2014). Lastly, participants were encouraged to use Autosummarizer (<http://autosummarizer.com/>), a website which uses AI to automatically produce summaries of texts.

These tools are relatively lightweight and require little to none installation of programs, so were far more readily amenable to the participants. A more complex tool was TXM, which stands for ‘textometry’, a methodology allowing quantitative and qualitative analysis of textual corpora, by combining developments in lexometric and statistical research with corpus technologies (<http://textometrie.ens-lyon.fr/?lang=en>) (Heiden, 2010). It allowed for a more granular analysis of language features, requiring the integration of a specific language model, the splitting of speakers, the conversion of data into computer readable XML language, and the lemmatization of the data.

4.2 Emotion recognition tools

Emotion recognition tools are often developed and used in the field of Social Signal Processing (SSP), where the goal is to investigate and develop machines that are socially intelligent; this implies that they are capable of recognizing and interpreting social and affective signals automatically (Vinciarelli et al., 2009). Tools exist that can extract characteristic speech parameters which can subsequently be used in machine learning software to find and learn new patterns (e.g., emotions) from this data. Applying these tools requires programming, so for social scientists and historians this is only feasible to apply in tandem with a computer scientist. What was key to convey to the participants, is the existence of this non-textual dimension of the recording. Emotion is indeed dealt with in the discipline ‘discourse analysis’, but not in a multimodal way. A way to profit from this tool, would be to identify particular emotions in an entire corpus, as a basis to go back to the interpretation of a single interview. Participants were first presented with the concept. During the session the use of the linguistic tool Praat (<http://www.fon.hum.uva.nl/praat/>) was demonstrated, showing the silences in a corpus and how these can be relevant for emotional expression analysis (Boersma, 2020). The hands-on component encouraged participants to get familiar with different speech features, for example, how to digitally detect and analyse these using a voice recording (Truong et al., 2013; Van den Heuvel & Oostdijk, 2016).

5 Pedagogical considerations for preparation and evaluation

Prior to the workshop, participants were invited to reflect on their own research trajectory and provide us with a short narrative of a typical research journey they had undertaken when working with interview data. Based on this homework, we assessed and visualized their workflows, and constructed a series of typical ‘research trajectory’ flow charts. This enabled us to come up with a high-level simplified trajectory and to identify how and where the digital tools might fit into the researchers’ workflow. This is illustrated in Figure 4.

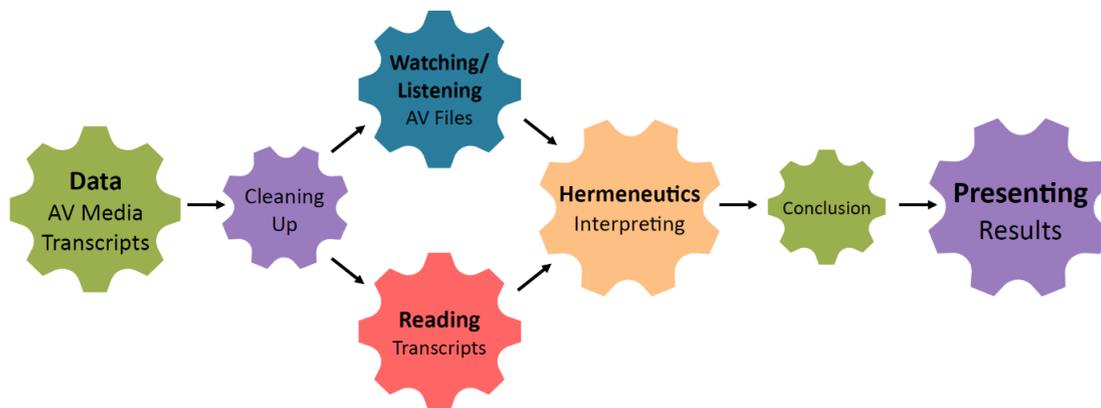


Figure 4: Poster showing typical research trajectory for scholars using interview data (cogs), annotated with specific tools that they worked on in the workshops.

Anticipating that the diversity of participants and tools would make the organization of the workshop complex, it was essential to follow principles in the design of the workflow that ensured ‘satisfying user and research experiences’. To this end we took great care of ensuring a basic level for preparation: we gathered information on the participants on their level of digital shrewdness; prepared data familiar to the participants in both a common language (English) and in their native language; assigned homework in order to make participants become familiar with the tools; and ensured that a participant with advanced digital skills was present in each of the language groups. After each session, each language group was asked questions on their experience and their assessment of tools’ strengths and weaknesses. We also asked them if they would use the tool in their own research and if so, in what way. Lastly, we asked them for suggestions that we could pass on to the tool’s developers.

6 Evaluation of tools

Even before the workshops started, some participants struggled to download unfamiliar software prior to meeting, suggesting a lack of basic technical proficiency. This can turn out to be a significant barrier to the use of open source tools that often require a bit more familiarity with, for example, laptop operating systems. It was useful to have speech technologists sitting amongst the scholars, witnessing first-hand some of the really basic challenges in getting started.

6.1 Evaluation of the Transcription Chain

The first hurdle with which the participants were confronted was the slow pace of the ASR process for particular languages. This meant that not all participants were able to fully experience working with the OH-portal. Those who did get the chance to play around with correcting the ASR results, appreciated the simple design and usability of the T-Chain. One participant remarked that they were glad the tool did not require much technical know-how. A happy surprise was that the software catered to all native languages represented by the different groups.

With regard to criticisms, lack of transparency seemed to be the overarching theme for all groups. During the long waiting time for results to show, participants were bothered by having no insight in what was happening behind the scenes. Several participants expected more information about the ASR-engines in terms of accuracy and speed. The interface in some cases was unclear, for example, there were two options to choose as ASR-engines (Dutch NL and Dutch OH), but the difference between the two was not clear. Similarly, the OCTRA-editor included for correcting ASR results had a lot of shortcuts, but these were unclear to those unfamiliar with speech technology terminology.

Attendees also expressed concern about the issue of privacy in sources that were uploaded to the web-based system, indicating that this may well be a problem in using the OH portal in their own research. Are the data stored somewhere after the processing is completed? They further questioned who owned the ASR-engines. Will the ASR results be used for improvement of the engine? One

British participant remarked that without answers to these questions, it would be hard to justify using the tool in front of an academic research ethics committee.

When asked for possible improvements on the T-Chain, participants stated that more documentation, on the technical, ethical and legal side, was needed. They pointed to the need for dejargonisation of terms or brands (like OCTRA) and further description or help added. One participant had the idea of a quick list on the side of the OCTRA-editor, where all the shortcuts to actions could be listed. File limits were of course also a problem. Participants would have liked to be able to upload not just wave files of a limited length, but also mp3 and different video files. More options for file conversion and export were requested, such as automatic conversion to TEI-encoded documents. Lastly, one feature that a British researcher would appreciate, is diarisation: the automatic detection of a speaker change, as well as a way of visualizing it.

To conclude, most of the participants were intrigued by the general concept of a T-Chain with its speech-to-text and alignment software. Almost all of them saw ASR as a way of potentially easing the transcription process. With some improvements, participants could see themselves using this resource for this purpose, although a long period of acculturation would be needed. Some doubted whether they would use the included OCTRA editor for correcting results, and stated they would feel 'more comfortable' using an external tool or word processor for this purpose.

6.2 Evaluation of text annotation tools

Overall, the familiarity of annotation across disciplines made both NVivo and ELAN accessible to participants. But the vastly different terminology and user interface meant that users had to spend additional time acquainting themselves to the tool's unique layout before being able to annotate. What would help is a uniformity of language and terminology for features that all tools have in common. A useful step to take is the dejargonising of the interface, or creating custom user interfaces for different types of users.

The NVivo tool worked particularly well with written transcripts, and allowed users to actually see mark-up and notes in the context of a transcript. Being able to collate all documents related to a single research project proved to be a clear benefit of the tool, with one user commenting that ELAN had a much more visual display and worked solely with audio and video data sources. But the learning curve of NVivo was steep for all participants unfamiliar with the software. Many experienced it as overwhelming and hard to work with. Another issue was how closed-off the package is. With barely any useful export and import functions, it is nearly impossible to use NVivo in combination with other software. ELAN, on the other hand, had good support for importing and exporting files.

Some users named working with ELAN as a 'pleasant experience', as it had a clear and comprehensible interface (the same on Mac and PC, which NVivo has not). The concept of 'tiered' annotations was found interesting, and a useful way of visualizing codes/notes. With ELAN, codes and annotations are tiered and placed on a timeline, something which NVivo lacks. While ELAN was taken up more quickly by the participants, its focus on the particularities of linguistics was experienced as a hurdle. Specifically, users disliked the lack of the possibility to make a distinction between transcriptions and annotations/codes in the tiers. It seemed that transcribing in ELAN is less suited for the in-depth interviews, typically used in Oral History, than briefer interviews that are often analysed by (socio-)linguists.

All in all, it was hard for participants to imagine getting out of their 'comfort zone'. The amount of time needed to become familiar with the features of the tools, and to actually experience the benefits was too short. Some expressed interest in exploring the tool further while others were turned off by the idea of using such intricate tools in general. The reason is that a choice of a particular annotation tool leads to an engrained practice of research that cannot be easily traded for an alternative. The most open to change were those who already used existing similar software such as ATLAS.ti. An experienced participant pointed to the suitability of ELAN for research where multimodal annotation/coding was required, with video/audio and text in tandem. The less digital savvy researchers suggested offering functionalities with different levels of complexity: a Simple Mode and an Expert Mode. This would make simple actions a lot more accessible to researchers with second to no experience with digital research.

6.3 Evaluation of on-the-fly linguistic analysis tools

Whereas the introduction to gain insight in the generic linguistic tools and their shortcomings/opportunities was very much welcomed, the hands-on components were met with varied reactions. Overall, the participants enjoyed these tools, and referred to them as “easy-to-use”, “simple”, and “lightweight”. *Voyant* and *Autosummarizer*, while not necessarily useful for drawing conclusions, inspired the participants to think about their own process of getting insight into a text and summarizing it. The limited amount of text that can be analysed was perceived as a barrier for the take up. There was an overall need to already be informed in a very concrete way about the added value for use of these tools by non-linguists. Exploring what the possibilities could be, proved to be demanding within the available time slot and only of interest to those interested in experimenting. What became clear is that sociolinguists may benefit from the use of the *Voyant* word frequency functionality.

Although the use of word frequency raises controversy within linguistics, it is widely accepted that frequent words may influence phonetic change, and also may act as ‘locus of style’ for a given speaker (Hay, Foulkes 2016, p. 324). At the same time, it seemed that *Voyant* was not sophisticated enough to process uncleaned transcriptions. Not everyone found the tools easy-to-use. *Voyant*, with its many different windows, was described by one researcher as “incredibly frustrating”, another participant wanted to know how the word clouds were generated. *Stanford NLP* a tool for more ‘heavy-duty’ linguistic analysis, was hard for the oral historians to relate to.

6.4 Evaluation of a textometry tool

Among advanced digital scholars, *TXM* was the most liked tool of all. One participant described it as the “most complete and transparent tool used so far”. The combination of both power and transparency gave the researchers lots of inspirations and ideas on how to use the software. Functionalities that were mostly appreciated were the *Tree Tagger*, the visualisation of concordances and co-occurrences, and the ease with which it was possible to get an overview of the complete corpus. Aside from the tool itself, the clear printouts of instructions in this session that each could follow at his or her own pace, were greatly appreciated. They formed a counterbalance to the complicated nature of *TXM*.

Interest was expressed for using co-occurrences and concordance in analysing both separate texts and sub-corpora. Some participants thought of this as a wholly new way of interacting with the data and discovering new aspects, while others thought it could prove or disprove ‘hunches’ they had about data, in that it is possible to quantitatively prove prevalence of certain words in text. A specific use for *TXM* that was mentioned was analysing differences in gender. *TXM* allows users to tag lets a user tag the gender of an interviewee and interviewer, which makes it possible to quantitatively analyse difference in the way these people speak, and making it a good tool to evaluate the entire corpus.

For most participants without experience with digital tools, the complexity of *TXM* was a hurdle. Some encountered problems with the terminology used in the program (e.g. what is a partition?). Several groups found the need to pre-process the data time-consuming and complicated. What could be improved of *TXM* is making the interface more approachable, adding more colour, turning the tool into a web-based service, and incorporating a glossary in which all the terms used in the software are explained. There were exceptions however: one user noted that “[*TXM* was] a bit of a struggle at first, but this helps you to do a close reading of an interview, and I think it fits perfectly within my traditional hermeneutical approach”.

Overall, it appeared difficult to understand how to attribute meaning to the frequency of a particular term in the entire corpus of interviews, when being used to focus on the interpretation of a single interview. *TXM* can offer insights in features of the interview process in its entirety, such as: the relation between words expressed by interviewer and interviewee, the difference in active and passive use of verbs between gender, age or profession, or the specificity of certain words for a respondent. In some ways, this might require the scholar to temporally disregard the individuality of the person talking, and switch from close listening to interpretations in which scale and numerical relations can be relevant. This requires a widening of methodological perspective in data analysis.

6.5 Evaluation of emotional recognition tools

The session on emotion recognition was the most remarkable one. While the technology behind it was the most complex, the relevance of identifying emotional features was clear to everyone. Participants first of all highly appreciated the outline of the discipline of Social Signal Processing (SSP). Even though Praat, as a tool for computational linguistics, was perceived as complex, the presentation illustrated its power in a way that all participants could easily relate to it. Here as well, the step by step guide provided was appreciated. One participant observed that she struggled less with terminology in this last session. She felt accomplished in having become familiar with technical terms like “tiers”. Still the step of connecting this dimension of the data to their own research practice was difficult to make. Experimenting with Praat was an enriching experience, but even after fully understanding how differences in pitch, speed and silences can be extracted from the data, they were doubtful on whether such applications could be integrated in social science or oral history research. How do you blend in insights with regard to scale, frequency and paralinguistic features into the classic interpretation of the interview data?

7 Conclusion

Interview data provide rich and promising information sources with which to engage in methodological interdisciplinary or multidisciplinary conversations. Our analysis of the user experience before, during and after the workshop suggests that scholars are open to cross-fertilization. At the same time, scholars are only willing to integrate a digital tool into their existing research practice and methodological mindset, if it can easily be used or adapted to their needs. The limited functionality of the free easy-to-use tools, and the observed methodological and technological complexity and jargon-laden nature of the dedicated downloadable tools, were both seen as significant barriers, despite the availability of clear documentation. Time investment is key to really grasp the essence of a tool. This means that addressing the right audience at the right stage of their career is crucial. It could also mean that we have to offer different trajectories with different levels of complexity to different types of scholars. With regard to assumptions we had about the potential of the tools to affect existing research practices - questioning the dogma of full verbatim transcription, exploring a collection with text mining tools, integrating the emotional features of an interview into an analysis - it is too early to be able to draw any conclusion. The workshops offered insights in new approaches to a broad group of researchers, but we have only scratched the surface of what is possible. A much longer and intensive engagement with data and tools in small multidisciplinary teams is necessary to test our assumptions. We intend to continue to publish the results of our endeavours through case studies, that can be selected from the multilingual archive of interviews on the topic of migration, that was created by the team for the workshops.

In terms of improving the take up among social scientists and historians who work with interview data, of what the CLARIN infrastructure has to offer, it is clear that jargon is an obstacle. More user friendly, well-documented and stable tools would be welcomed, especially in a form that makes it possible to skip the lengthy process of installing software. More specifically, it may be worth exploring the integration of open source annotation tools into the latter end of the T-Chain to enable a seamless experience, and thereby moving towards the ever-appealing concept of the workbench. Moreover, as in all ASR services, privacy concerns must be addressed so that users have a very clear understanding of what will happen to a source once uploaded to a tool. In this respect, an explicit GDPR-compliant data processing agreement may allay worries. A closer collaboration with the CLARIN Ethical and Legal Committee (CLIC) is therefore recommended. As different disciplines and tools for interview data use varying metadata schema, work needs to be done on mapping and crosswalks and export of marked up formats. Empowering the curators and publishers of interview data to get their collections ‘analysis- and tool ready’ is useful, as is encouraging them to also make use of text mining tools, such a term extraction, for enhancing resource discovery.

References

- Armaselu, F., Danescu, E., Klein, F. 2019. Oral History and Linguistic Analysis. A Study in Digital and Contemporary European History. *Linköping Electronic Conference Proceedings* (159), 13.
- Boersma, P., Weenink, D. 2020. Praat: doing phonetics by computer [Computer program]. Version 6.1.10, retrieved 23 March 2020 from <http://www.praat.org/>.
- Boyd, D. (ed). 2013. Oral History in the Digital Age. *The Oral History Review*, 40(1): i-iii doi:10.1093/ohr/oht038.
- Corti, L., Gregory, A. 2011. CAQDAS Comparability. What about CAQDAS Data Exchange? [42 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 12(1), Art. 35, <http://nbn-resolving.de/urn:nbn:de:0114-fqs1101352>.
- Corti, L., Fielding, N. 2016. *Opportunities From the Digital Revolution: Implications for Researching, Publishing, and Consuming Qualitative Research*. SAGE Open. <https://doi.org/10.1177/2158244016678912>.
- De Jong, F.M.G., van Hessen, A., Petrovic T., Scagliola S. 2014. Croatian Memories: speech, meaning and emotions in a collection of interviews on experiences of war and trauma. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, <http://www.lrec-conf.org/proceedings/lrec2014/index.html>.
- De Jong, F.M.G., Ordelman, R.J.F., Scagliola S. 2011. Audio-visual Collections and the User Needs of Scholars in the Humanities: a Case for Co-Development. *Proceedings of the 2nd Conference on Supporting Digital Humanities (SDH 2011)*.
- Freund, A. 2009. Oral history as process-generated data, *Historical Social Research*, 34 (1): 22–48.
- Heiden, S. 2010. The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme. In K. I. Ryo Otoguro (Ed.) *24th Pacific Asia Conference on Language, Information and Computation*: 389-398. Institute for Digital Enhancement of Cognitive Development, Waseda University. Online: <https://halshs.archives-ouvertes.fr/halshs-00549764>.
- Hay, J., Foulke, P. 2016. The evolution of medial /t/ over real and remembered time. *Language*, 2016, 92: 298-330.
- Kisler, T., Schiel, F., Sloetjes, H. 2012. Signal Processing via Web Services: the Use Case WebMAUS. *Proceedings of DH2012*, 30-34, 2012, Hamburg.
- Pömp, J., Draxler, Chr. 2017. OCTRA – A Configurable Browser-Based Editor for Orthographic Transcription. *Tagungsband Der 13. Tagung Phonetik Und Phonologie Im Deutschsprachigen Raum*, 145-148. Berlin.
- Portelli, A. 2006. What makes oral history different, in: R. Perks and A. Thomson, *The Oral History Reader* (New York) 36.
- Sinclair, S., Rockwell G., 2016. *Voyant Tools*. Web. <http://voyant-tools.org/>.
- Truong, K. P., Westerhof, G. J., Lamers, S.M.A., de Jong, F.M.G., Sools, A. 2013. Emotional expression in oral history narratives: comparing results of automated verbal and nonverbal analyses. *Proceedings of the Workshop on Computational Models of Narrative CMN 2013 Hamburg, Germany*.
- Van den Berg, H., Scagliola, S., Wester, F. (eds) 2010. *Wat veteranen vertellen; verschillende perspectieven op verhalen over ervaringen tijdens militaire operaties (What veterans tell us; different perspectives on biographical interviews about experiences during military operations)*, Pallas Publications. <http://www.watveteranenvertellen.nl/>.
- Van den Heuvel, H., Draxler, C., Van Hessen, A., Corti, L., Scagliola, S., Calamai, S., Karrouche, N. 2019. A

Transcription Portal for Oral History Research and Beyond. *Proceedings DH2019, Utrecht, 10-12 July 2019*. <https://dev.clariah.nl/files/dh2019/boa/0854.html>.

Van den Heuvel H., Oostdijk, N.H.J. 2016. Falling silent, lost for words ... Tracing personal involvement in interviews with Dutch war veterans. *Proceedings LREC2016, Portorož, Slovenia*: 998-1001, http://www.lrec-conf.org/proceedings/lrec2016/pdf/104_Paper.pdf.

Van den Heuvel, H., van Hessen, A., Scagliola, S., Draxler, C. 2017. Transcribing Oral History Audio Recordings – the Transcription Chain Workflow. *Poster at EU. Clarin Conference, Budapest, September 18/19- 2017. EU. Clarin Conference*.

Vinciarelli, A., Pantic, M., Bourlard, H. 2009. Social signal processing: Survey of an emerging domain. *Image and vision computing*, 27(12): 1743-1759.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. 2006. ELAN: a professional framework for multimodality research. *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*. 1556-1559.